FITTING THE VARIANCE-GAMMA MODEL TO FINANCIAL DATA

EUGENE SENETA

Abstract

This paper has as its main theme the fitting in practice of the variance-gamma distribution, which allows for skewness, by moment methods. This fitting procedure allows for possible dependence of increments in log returns, while retaining their stationarity. It is intended as a step in a partial synthesis of some ideas of Madan, Carr and Chang (1998) and of Heyde (1999). Standard estimation and hypothesis-testing theory depends on a large sample of observations which are independently as well as identically distributed and consequently may give inappropriate conclusions in the presence of dependence.

Keywords: Variance-gamma; log returns; subordinator model; skewness; increments; dependence; stationarity; method of moments; estimation; martingale; characteristic function; t-distribution; 'Student' processes

2000 Mathematics Subject Classification: Primary 60G15
Secondary 60E07; 60E10; 62M05

1. Introduction and setting

The general model we consider gives the price P_t of a risky asset over time $t \geq 0$ by

$$P_t = P_0 \exp\{ct + \theta T_t + \sigma W(T_t)\}, \tag{1}$$

where c, θ, and σ (> 0) are real constants. The (market) activity time, $\{T_t\}$, is a positive increasing random process with stationary differences $\tau_t = T_t - T_{t-1}$ for $t \geq 1$, which is independent of the standard Brownian motion $\{W(t)\}$. The corresponding log returns at unit time intervals are then given by

$$X_t = \log P_t - \log P_{t-1} = c + \theta(T_t - T_{t-1}) + \sigma(W(T_t) - W(T_{t-1})). \tag{2}$$

We shall assume that $\mathrm{E}\,\tau_t < \infty$, and so without loss of generality that

$$\mathrm{E}\,\tau_t = 1 \tag{3}$$

to make the expected activity time change over unit calendar time equal to one unit, the scaling change in time being absorbed into θ and σ, noting that

$$\sigma(W(T_t) - W(T_{t-1})) \stackrel{\mathrm{D}}{=} \sigma(T_t - T_{t-1})^{1/2} W(1). \tag{4}$$

It follows that

$$\mu := \mathrm{E}\,X_t = c + \theta, \tag{5}$$

$$\mathrm{var}(\sigma(W(T_t) - W(T_{t-1}))) = \sigma^2. \tag{6}$$

Using the fact that, conditional on T_t and T_{t-1}, $X_t - \mu \sim \mathcal{N}(\theta(\tau_t - 1), \sigma^2\tau_t)$, it follows that

$$\mathrm{E}((X_t - \mathrm{E}\,X_t)^2) = \sigma^2 + \theta^2 M_2, \tag{7}$$

$$\mathrm{E}((X_t - \mathrm{E}\,X_t)^3) = 3\theta\sigma^2 M_2 + \theta^3 M_3, \tag{8}$$

$$\mathrm{E}((X_t - \mathrm{E}\,X_t)^4) = 3\sigma^4(1 + M_2) + 6\sigma^2\theta^2(M_2 + M_3) + \theta^4 M_4, \tag{9}$$

where $M_i = \mathrm{E}((\tau_t - 1)^i)$, $i = 2, 3, 4$, recalling (3) and assuming finiteness of $\mathrm{E}\,\tau_t^i$ for $i = 2, 3, 4$ as necessary.

The coefficients of skewness (a measure of asymmetry) and of kurtosis (whose size in excess of 3 measures the degree that tails are fatter than the normal) are defined respectively by

$$\beta = \frac{\mathrm{E}((X_t - \mathrm{E}\,X_t)^3)}{(\mathrm{E}((X_t - \mathrm{E}\,X_t)^2))^{3/2}} \tag{10}$$

and

$$\kappa = \frac{\mathrm{E}((X_t - \mathrm{E}\,X_t)^4)}{(\mathrm{E}((X_t - \mathrm{E}\,X_t)^2))^2}. \tag{11}$$

The case $T_t = t$ of the model (1) is the classical geometric Brownian motion (GBM) model, with corresponding incremental log returns being independent and identically distributed (i.i.d.) Gaussian. When the increments $\{\tau_t\}$ in the model (1) are *independent* as well as stationary, the model was treated by Hurst *et al.* (1997, pp. 100–102), whence (7)–(9) also follow (the independence of increments assumption does not affect their formulae (17)–(20)).

Setting the parameter $\theta = 0$ in (2) shows, using (3) and (5), that X_t is distributed *symmetrically* about its mean $\mu = c$ with $\mathrm{var}\,X_t = \sigma^2$ from (7). From (9) and (7) the kurtosis is then $3(1 + M_2)$. Noting that symmetry is in practice usually the case for financial data, Hurst *et al.* (1997, p. 103 ff.) give a thorough survey of (subordinator) models under $\theta = 0$ and the additional assumption of independence of increments.

The model (1) with arbitrary θ, independent increments, and very specific (gamma) distribution of τ_t was very fully developed by Madan *et al.* (1998); see also the earlier Madan and Milne (1991). The original variance-gamma (VG) model (Madan and Seneta (1990)) is the restricted case where $\theta = 0$, and is called the 'symmetric variance-gamma' model by Madan *et al.* (1998), the 'log variance gamma model' by Hurst *et al.* (1997), and 'The Madan and Seneta log variance gamma model' by Hurst and Platen (1997).

Under this additional assumption of independence of increments in model (1), $\{X_t\}$, $t \geq 1$, is a sequence of i.i.d. random variables, so classical statistical procedures for estimation and hypothesis testing (such as maximum likelihood) are readily applied in specific cases where the probability density function (PDF) of X_t is available in computationally tractable form (see Hurst *et al.* (1997)). Notice that our discussion of estimation only requires the model to be discrete, i.e. that (2) hold for $t = 1, 2, \ldots$.

It is appropriate to note on this occasion that the papers authored by Hurst, Platen and Rachev cited were written in Chris Heyde's Stochastic Analysis Group at the Australian National University, and are closely related to some of his own work to which we now come.

2. The FATGBM subordinator model

To account for the facts that typical log-returns data show higher peaks and heavier tails than the Gaussian, and evidence of strong and persistent dependence of absolute values and of squares of log returns, Heyde (1999) developed the fractal activity time geometric Brownian motion (FATGBM) model, which can be described by taking $\theta = 0$ in (1) and (2) and taking the process $\{T_t\}$ so that its increments $\{\tau_t\}$ exhibit long-range dependence (LRD) and have heavy tails (as well as, consequently, symmetry, with $\mathrm{E}\,X_t = c$, $\operatorname{var} X_t = \sigma^2$). There is strong empirical evidence (Heyde and Liu (2001)) that the process $\{T_t - t\}$ may be taken as self-similar of index H, $\frac{1}{2} < H < 1$. That is, for positive c and $t \geq 0$

$$T_{ct} - ct \overset{\mathrm{D}}{=} c^H (T_t - t),$$

in which case the desired properties of τ_t follow.

The statistical estimation procedure employed is simply the method of moments, relying on the stationarity of increments, consistency following from the ergodic theorem. Recalling that we are considering the case $\theta = 0$, the parameters $\mathrm{E}\,X_t = c$ and $\operatorname{var} X_t = \sigma^2$ in the model of (1) and (2) may be consistently estimated in the usual way from the long observed sequence $\{X_t\}$, $t = 1, \dots, n$ by

$$\hat{c} = \bar{X} = \frac{\sum_{t=1}^{n} X_t}{n}, \qquad \hat{\sigma}^2 = \frac{\sum_{t=1}^{n} (X_t - \bar{X})^2}{n}.$$

For the further specific estimation to do with the process $\{T_t\}$ of the FATBGM model, e.g. of H, we refer the reader to Heyde (1999) and Heyde and Liu (2001).

Potential *asymmetry* in the distribution of X_t can be allowed for in the FATBGM model at the cost of an additional parameter via (1), as is easily seen from (8). Heyde and Gay (2000) take $\theta = -\sigma^2/2$, as a result of modelling P_t by the stochastic difference equation:

$$\mathrm{d}P_t = P_t\{c\,\mathrm{d}t + \sigma\,\mathrm{d}W(T_t)\}.$$

The skewness and kurtosis coefficients (10) and (11), when they exist, can be consistently estimated for the general model of (1) and (2) by

$$\hat{\beta} = \frac{\sum_{t=1}^{n} (X_t - \bar{X})^3}{\left(\sum_{t=1}^{n} (X_t - \bar{X})^2\right)^{3/2}}, \tag{12}$$

$$\hat{\kappa} = \frac{\sum_{t=1}^{n} (X_t - \bar{X})^4}{\left(\sum_{t=1}^{n} (X_t - \bar{X})^2\right)^{2}}, \tag{13}$$

where

$$\bar{X} = \frac{\sum_{t=1}^{n} X_t}{n} = \hat{\mu}. \tag{14}$$

The FATGBM model appears to be simple, practical and remarkably successful in capturing the movements of a wide range of data, and reflects probabilistic insight and breadth of technique. The parameters c and σ^2 figure as universal, independent of the process $\{\tau_t\}$.

My own interest up to the appearance of Madan and Seneta (1990) had been in the symmetric VG model. Dilip Madan had moved to the University of Maryland at College Park from the University of Sydney, and I was visiting the University of Virginia when we put the last touches to that paper in 1989. After some years of my own inactivity in this area, I became aware of Chris Heyde's penetration of yet another probabilistic field, that of financial modelling, and asked him to speak at the University of Sydney on 26 March 1999, on what was soon to be published as Heyde (1999). His opinion of the symmetric VG model was that, although it was heavier tailed than the normal, the tails were not heavy enough to account for what were occasionally relatively extreme values of X_t; and of course the increments in the VG model had been assumed independent. Although the underlying stochastic process $\{\log(P_t/P_0)\}$, $t \geq 0$, was closed under convolution and analytically convenient, a treatment such as he proposed could effectively reconstruct the $\{T_t\}$ process numerically from data. My occasional visits to

Canberra allowed me to learn more about the FATGBM ideology, and to obtain advice from Chris on the statistical fitting of the symmetric VG model. The VG model had continued to be of interest. I had had several inquiries as to how to fit it to real data, and I needed to supervise the fourth-year Honours project of Annelies Tjetjep (2002). My focus was Dilip Madan's post-1990 successful extension and application of the VG models (see items in the references co-authored by Madan, especially Madan *et al.* (1998)), where fitting from data as well as modelling are integral issues. Madan *et al.* (1998), in the guise of its Research Report predecessor, was already described in a monograph (Epps (2000)).

Dilip Madan, Wake Epps, and Eckhard Platen very kindly supplied me with current materials and information, as of course did Chris Heyde.

The next section considers the procedure and effect of fitting the (general) VG by allowing for *dependence of increments* while retaining their stationarity. We do not address specifically the issue of adequacy of tail structure of the VG distribution. Important new work by Heyde and Kou (2002) suggests, in any case, that heavy-tail (power-law) structure is not easily distinguishable in practice from exponential-tail structure.

3. Fitting the variance-gamma model

The PDF of X_t defined by (2) is given for the general VG model by

$$f_X(x) = \frac{2\exp(\theta(x-c)/\sigma^2)}{\sigma\sqrt{2\pi}\,\nu^{1/\nu}\Gamma(1/\nu)}\left(\frac{|x-c|}{\sqrt{2\theta^2/\nu+\sigma^2}}\right)^{1/\nu-1/2} K_{1/\nu-1/2}\left(\frac{|x-c|\sqrt{2\sigma^2/\nu+\theta^2}}{\sigma^2}\right) \tag{15}$$

for $-\infty < x < \infty$ (Madan *et al.* (1998)), where $\nu > 0$ and the other parameters are as described previously. Here $K_\eta(\cdot)$ is a modified Bessel function of the third kind (Erdélyi *et al.* (1953); Hurst *et al.* (1997, p. 107)) with index η, given for $\omega > 0$ by

$$\begin{aligned} K_\eta(\omega) &= \frac{1}{2}\int_0^\infty \exp\left\{-\frac{\omega}{2}(v^{-1}+v)\right\} v^{\eta-1}\,\mathrm{d}v \\ &= \int_0^\infty \exp(-\omega\cosh t)\cosh(\eta t)\,\mathrm{d}t. \end{aligned}$$

There is ambiguity in terminology (Tjetjep (2002)): Madan and Seneta (1990) and Madan *et al.* (1998) describe $K_\eta(\cdot)$ as a modified Bessel function of the second kind. However, Whittaker and Watson (1915, p. 367) describe $\cos(\eta\pi)K_\eta(\omega)$ as a modified Bessel function of the second kind. Since in the setting above $\eta = 1/\nu - \frac{1}{2}$ involves a parameter ν to be estimated, care has to be taken in computing if we use the functional form of $f_X(x)$, for example in maximum likelihood estimation, as may be done with the package MATLAB® which permits use of the function $K_\eta(\cdot)$.

The characteristic function of X_t (Madan *et al.* (1998, p. 83)) is

$$\phi_X(u) = \mathrm{e}^{\mathrm{i}cu}(1 - \mathrm{i}\theta\nu u + (\sigma^2\nu u^2)/2)^{-1/\nu} \tag{16}$$

for $-\infty < u < \infty$, since the distribution of $\tau_t = T_t - T_{t-1}$ is given (in the setting of our Section 1) by the gamma PDF:

$$f_\tau(w) = \begin{cases} \dfrac{1}{\nu^{1/\nu}}\dfrac{w^{1/\nu-1}\mathrm{e}^{-w/\nu}}{\Gamma(1/\nu)}, & w > 0, \\ 0 & \text{elsewhere,} \end{cases} \tag{17}$$

so that $\mathrm{E}\,\tau_t = 1$, as required, and $M_2 = \operatorname{var}\tau_t = \nu$.

Consequently, after computing M_3 and M_4 and using (7)–(9), we finally obtain the following (Madan *et al.* (1998, p. 85)), where $\mu = \mathrm{E}\,X_t$:

$$\mathrm{E}\,X_t = c + \theta, \tag{18}$$

$$\operatorname{var} X_t = \sigma^2 + \theta^2\nu, \tag{19}$$

$$\mathrm{E}((X_t - \mathrm{E}\,X_t)^3) = 2\theta^3\nu^2 + 3\sigma^2\theta\nu, \tag{20}$$

$$\mathrm{E}((X_t - \mathrm{E}\,X_t)^4) = 3\sigma^4\nu + 12\sigma^2\theta^2\nu^2 + 6\theta^4\nu^3 + 3\sigma^4 + 6\sigma^2\theta^2\nu + 3\theta^4\nu^2. \tag{21}$$

The case $\theta = 0$ (Madan and Seneta (1990)) is the original (symmetric) variance-gamma distribution, with the distribution of $X_t - \mu$ (here $\mu = c$) consequently having the simple real-valued characteristic function

$$\phi_X(u) = \left(1 + \frac{\sigma^2\nu u^2}{2}\right)^{-1/\nu}. \tag{22}$$

Letting $\nu \to 0+$ results in $\phi_X(u) = \exp(-\sigma^2 u^2/2)$, so X has an $\mathcal{N}(0, \sigma^2)$ distribution.

In the general case (15), the moment estimators (14), $\widehat{\operatorname{var}} X_t = \sum_{i=1}^n (X_i - \bar{X})^2/n$, and (12) and (13) may be used to estimate μ, $\operatorname{var} X_t$, β, and κ (given by (10) and (11)), whence moment estimates of c, θ, σ^2, and ν may be obtained by solving (18)–(21). The solutions may in practice require iterative procedures, but the following simple arguments can be used to begin such a procedure, and *will sometimes suffice for estimation*, as we will demonstrate in several examples.

First note that

$$\beta = \frac{2\theta^3\nu^2 + 3\sigma^2\theta\nu}{(\theta^2\nu + \sigma^2)^{3/2}} \qquad \text{(skewness)},$$

$$\kappa = 3 + \frac{3\sigma^4\nu + 12\sigma^2\theta^2\nu^2 + 6\theta^4\nu^3}{(\theta^2\nu + \sigma^2)^2} \qquad \text{(kurtosis)}$$

from (19)–(21), so ignoring terms in θ^2, θ^3, θ^4, we have

$$\mu = c + \theta, \tag{23}$$

$$\operatorname{var} X_t = \sigma^2, \tag{24}$$

$$\beta = \frac{3\theta\nu}{\sigma}, \tag{25}$$

$$\kappa = 3(1 + \nu). \tag{26}$$

Thus, from $\widehat{\operatorname{var}} X_t$, $\hat{\kappa}$, $\hat{\beta}$, and $\hat{\mu}$ in turn we may successfully obtain approximations to $\hat{\sigma}^2$, $\hat{\nu}$, $\hat{\theta}$, and $\hat{c}$.

If $\hat{\theta}$ is small, then the full equations (18)–(21) will be close-to-satisfied by $\hat{\sigma^2}$, $\hat{\nu}$, $\hat{\theta}$, and $\hat{c}$ obtained from the approximate equations (23)–(26); hence these themselves form suitable (moment) estimates. There is no need for the independence of increments assumption here.

Example 1. (*Standard and Poor's 500 Index, January 1977–December 1981.*) The number of log daily price differences was $n = 1262$. We have

$$\hat{\mu} = 1.0752 \times 10^{-4},$$

$$\widehat{\operatorname{var}} X_t = 6.4474 \times 10^{-5},$$

$$\hat{\beta} = -0.023\,813,$$

$$\hat{\kappa} = 4.2661.$$

Then, from (23)–(26),

$$\hat{\sigma}^2 = 6.447 \times 10^{-5},$$
$$\hat{\nu} = 0.4220,$$
$$\hat{\theta} = -1.510 \times 10^{-4},$$
$$\hat{c} = 2.585 \times 10^{-4},$$

so these can in fact be taken as the moment estimators, since terms in $\hat{\theta}^2$, $\hat{\theta}^3$, $\hat{\theta}^4$ are relatively negligible.

It is plausible on account of the sizes of $\hat{\beta}$ and $\hat{\nu}$ to suppose that $\theta = 0$, so the symmetric model is plausible. Then the moment estimates for σ^2 and ν are unchanged, but $\hat{c} = (\hat{\mu} =)\ 1.0752 \times 10^{-4}$ is considerably affected for this numerical example, which casts into some doubt the assumption $\theta = 0$.

When the additional initial assumption of independence of increments is made, moment estimators are generally used as a starting point for maximum likelihood (ML) estimation. To reduce the number of parameters to be estimated by ML from four, the (moment) mean estimate of μ, $\hat{\mu} = \bar{x}$, is subtracted from the readings, giving the mean-corrected readings $\{z_i\}$, $i = 1, \dots, n$, where $z_i = x_i - \bar{x}$. These are regarded as a random sample (of i.i.d. readings) from a distribution (of $Z = X_t - \mu$) with PDF:

$$f_Z(z) = \frac{2\exp(\theta(z+\theta)/\sigma^2)}{\sigma\sqrt{2\pi}\,\nu^{1/\nu}\Gamma(1/\nu)}\left(\frac{|z+\theta|}{\sqrt{2\sigma^2/\nu+\theta^2}}\right)^{1/\nu-1/2} K_{1/\nu-1/2}\left(\frac{|z+\theta|\sqrt{2\sigma^2/\nu+\theta^2}}{\sigma^2}\right). \tag{27}$$

Analysis of the data in the above example gave (Tjetjep (2002)) the ML estimates $\hat{\nu}_{\mathrm{ML}} = 0.382$, $\hat{\sigma}_{\mathrm{ML}} = 8.03 \times 10^{-3}$, $\hat{\theta}_{\mathrm{ML}} = -8.31 \times 10^{-4}$.

When the original data x_i, $i = 1, \dots, n$ were used with the PDF (15) with parameter μ reintroduced, that is, with c in (15) replaced by $\mu - \theta$, attempts to find the values of $(\mu, \nu, \sigma, \theta)$ to maximize the likelihood function were unsuccessful, since the likelihood surface in the vicinity of the moment estimators appeared to be flat. However, with $\theta = 0$, the ML estimators in this presumed symmetric case were

$$\hat{\mu}_{\mathrm{MLS}} = 1.9752 \times 10^{-4}, \qquad \hat{\nu}_{\mathrm{MLS}} = 0.372, \qquad \hat{\sigma}_{\mathrm{MLS}} = 8.02 \times 10^{-3}.$$

The difference between $\hat{\mu} = 1.0752 \times 10^{-4}$ and $\hat{\mu}_{\mathrm{MLS}} = 1.9752 \times 10^{-4}$ may be due to the assumption of independence of increments underlying ML estimation, when this assumption is not appropriate. The plot of the sample autocorrelation function of the squared returns data of the readings shows the same evidence of LRD as in Heyde and Liu's (2001) data analysis of daily Standard and Poor's observations (November 1964–November 1999).

While the method of moments is lacking somewhat in precision, it compensates in terms of robustness.

4. Further examples, skewness, and goodness of fit

Example 2. In Standard and Poor's data (January 1992–September 1994; 691 days) analysed by Madan *et al.* (1998) the moment estimates reported (p. 90, footnote 16) were $\hat{\sigma} = 0.006233$, $\hat{\beta} = -0.2105$, $\hat{\kappa} = 7.8463$, so that, from (24)–(26), $\hat{\sigma}^2 = 3.885 \times 10^{-5}$, $\hat{\nu} = 1.615$, $\hat{\theta} = -2.707 \times 10^{-4}$ and certainly these can serve as reliable moment estimators. After ML fitting, the daily ν-estimate is reported as $\hat{\nu}_{\mathrm{ML}} = 365 \times 0.002 = 0.73$ for both the VG and symmetric VG (which is regarded as being an adequate model, since $\hat{\theta}_{\mathrm{ML}}$ is insignificant). Note the substantial decrease in the value $\hat{\nu}_{\mathrm{ML}}$ relative to $\hat{\nu} = 1.615$. The value of $\hat{\mu} = \bar{x}$ is not reported by Madan *et al.* (1998) and we do not know the precise methodology (e.g. mean-corrected data?) used by Madan *et al.* (1998) to obtain the ML estimates.

Example 3. Hurst *et al.* (1997, p. 115) report carefully obtained moment estimates for $n = 5198$ readings on X_t for the period 4 January 1973–30 July 1993 (omitting the large negative increment of 19 October 1987 due to the stock market crash). The daily readings were the Dow-Jones Industrial Average. The moment estimates reported were $\hat{\mu} = 3.0328 \times 10^{-4}$, $\widehat{\mathrm{var}} X_t = 8.1111 \times 10^{-5}$, $\hat{\beta} = 0.20080$, $\hat{\kappa} = 9.8141$.

Thus, from (23)–(26),

$$\hat{\sigma}^2 = 8.111 \times 10^{-5}, \tag{28}$$

$$\hat{\nu} = 2.271, \tag{29}$$

$$\hat{\theta} = 2.654 \times 10^{-4}, \tag{30}$$

$$\hat{c} = 3.788 \times 10^{-5}, \tag{31}$$

and again these may be taken as moment estimators.

The corresponding ML estimates for the *symmetric* VG (Hurst *et al.* (1997, p. 121, Table IV)) are

$$\hat{\mu}_{\mathrm{MLS}} = 1.5272 \times 10^{-4},$$
$$\hat{\sigma}^2_{\mathrm{MLS}} = 7.399 \times 10^{-5},$$
$$\hat{\nu}_{\mathrm{MLS}} = 0.5844.$$

Since at $\theta = 0$ in the general VG model, $\hat{\sigma}^2 = 8.111 \times 10^{-5}$ and $\hat{\nu} = 2.271$, the change under ML estimation is dramatic, and—although there are other possible explanations—the assumption of independent increments which validates the ML procedures may be responsible and the skew (here positive) may well be significant.

Finally, we note that in all three of the examples, the readings have been daily (so in model (2) unit time is 1 day). Skewness may be significant for such high-frequency readings. For lower frequency readings (where the unit of time is a week or a longer period) the log returns are close to symmetric because of aggregation. The model allowing for skewness, however, certainly seems appropriate for option-pricing data at weekly intervals (Madan *et al.* (1998); see also Section 5 below and Heyde and Gay (2000)).

4.1. Testing for goodness-of-fit

The testing procedures for goodness-of-fit such as the elementary chi-square goodness-of-fit procedure and the Kolmogorov and Anderson–Darling tests, or the likelihood ratio comparison procedures used by Madan *et al.* (1998) Hurst *et al.* (1997) and Hurst and Platen (1997), depend on the observations on X_t being i.i.d. (i.e. on independence as well as stationarity of increments) and the results produced may therefore be unreliable in the presence of dependence. The early statistical investigations of the (symmetric) VG model focused on the simplicity of the characteristic function

$$\phi_1(u) = \left\{ \frac{\nu}{m} \left(\frac{m}{\nu} + \frac{u^2}{2} \right) \right\}^{-m^2/\nu} \tag{32}$$

(which is misprinted in Madan and Seneta (1987b), (1987c)) to apply a characteristic function estimation method corresponding to transformed maximum likelihood, and relied on the assumption of i.i.d. increments of log-price. The parametrization leading to (32) was different from that in the present paper. If we denote estimates of the parameters in (32) by $\tilde{m}$ and $\tilde{\nu}$, then estimates $\hat{\sigma}^2$ of σ^2 and $\hat{\nu}$ of ν in our present parametrization leading to (22) would be given by

$$\hat{\sigma}^2 = \tilde{m}, \qquad \hat{\nu} = \frac{\tilde{\nu}}{\tilde{m}^2}.$$

In Madan and Seneta (1987b), the chi-square goodness-of-fit test indicated satisfactory fit of the symmetric VG model in 12 of 19 cases of large stocks on the Sydney Stock Exchange. The result may possibly have been improved by using the Anderson–Darling test or by allowing for skewness of the VG distribution, but it is also plausible that some of the bad fit was due to the dependence of increments.

It is probably well known that the chi-square goodness-of-fit test, in the face of positively associated identically distributed readings, will tend to reject the null hypothesis when it is true. In any case, here is a simple argument in this direction: suppose that p_i is the probability that a reading from a certain distribution falls into the ith region of a partition of sample space into k regions. With perfect positive association, a sample of n readings amounts to n identical numerical values. The expected number of values for the ith region is np_i, and so the value of the goodness-of-fit statistic is

$$\frac{(n - np_i)^2}{np_i}$$

for some i, with probability p_i. Thus the expected value of the goodness-of-fit statistic is

$$\begin{aligned}\sum_{i=1}^{k} \frac{(n - np_i)^2}{np_i} p_i &= n \sum_{i=1}^{k} (1 - p_i)^2 \\ &\geq n \sum_{i=1}^{k} \left(1 - \frac{1}{k}\right)^2 \\ &= n \left(\frac{k-1}{k}\right)(k-1),\end{aligned}$$

whereas under independence of the sample values, the expected value is (for large n) $k - 1$, as is well known.

Finally, the reader interested in characteristic function estimation methods using transformed maximum likelihood and i.i.d. readings may wish to consult the extension Madan and Seneta (1989) of Madan and Seneta (1987b), and corresponding numerical application of the methodology in Madan and Seneta (1987a), including the successful application to asymmetric stable laws.

5. Martingales and Markovian dependence

Madan *et al.* (1998) actually considers the model of (1) and (2) with the additional assumption of independence of VG-distributed increments, but with the overall mean μ of X_t defined by

$$\mu = m + \frac{1}{\nu}\left(1 - \theta\nu - \frac{\sigma^2\nu}{2}\right) + \theta, \tag{33}$$

where m is a constant (unrelated to m in (32)). For this to make sense it is necessary that

$$\theta + \frac{\sigma^2}{2} < \frac{1}{\nu}. \tag{34}$$

The reason for the parametrization (33) seems to be that the process $\{Y_t\},\ t \geq 0$, where

$$Y_t = Y_0 \exp\left\{\theta T_t + \sigma W(T_t) + \frac{t}{\nu}\ln\left(1 - \theta\nu - \frac{\sigma^2\nu}{2}\right)\right\} \tag{35}$$

is in fact a martingale. Thus, the no-arbitrage requirement holds, and the risk-neutral price of a European call option can be determined from the process $\{e^{rt}Y_t\}$ where r is a known, constant, continuously compounded interest rate. That is, in (33) m is taken as r and is not a parameter. In fact Madan *et al.* (1998, p. 88) obtained a closed-form expression for this price. The parameters σ^2, ν, and θ are determined from option prices of (weekly) observations on $\{e^{rt}Y_t\},\ t \geq 0$.

One conclusion reached by Madan *et al.* (1998) is that there is negative skewness present ($\hat{\theta}_{\mathrm{ML}} = -0.1436$ is substantial), which motivated our own interest in skewness of the preceding sections.

The structure of the martingale (35) can be seen to follow from the fact that the moment generating function (MGF) of X_t given by (2) (compare (16)) is

$$M_X(s) = \mathrm{E}(\exp\{sX_t\}) = \left(1 - \nu\left(\theta s + \frac{\sigma^2 s^2}{2}\right)\right)^{-1/\nu},$$

so, if (34) holds, $M_X(1)$ is well defined and

$$\mathrm{E}(\exp\{-\ln M_X(1) + X_t\}) = 1.$$

This device for constructing a GBM-type martingale $\{Y_t\}$ from the model (1) works providing the investments are assumed (additionally) to be independent and $M_X(1)$ is assumed well defined. We do not need the specific VG form, but the device does not work for the recently popular t-distribution (see Section 6 below) because the MGF does not exist for any $s \neq 0$.

The martingale device may be retained in a *discrete-time* version of the model of (1) and (2), where the dependence structure of $\{\tau_t\},\ t = 1, 2, \ldots$, is Markovian, and the process is in stationary regime, with the stationary distribution being (say) VG. The construction of one possible Markovian transition structure is well known from the Hastings–Metropolis algorithm of Markov chain Monte Carlo methods. This will define the process $\{X_t\},\ t = 0, 1, \ldots$, analytically. We do not pursue details here, but note that the moment estimation methods we have discussed are clearly applicable under such a discrete dependence model. Estimation of μ, σ^2, ν, and θ will lead via (33) to an estimate of m if required; that is, if we are statistically modelling the real-world movement of the price of a financial entity.

If m is prespecified as *known* r and actual option prices are used for $\{e^{rt}Y_t\}$, then the four moment-estimate equations (33) (with r in place of m), (24), (25), and (26) will overdetermine the three estimates $\hat{\sigma}^2$, $\hat{\nu}$, $\hat{\theta}$. If (24)–(26) alone are used, agreement will indicate $\bar{x}$, and $\hat{\mu}$ calculated from (33) will indicate roughly degree of model fit to option prices.

The condition (34) holds (if σ^2, ν and θ are replaced by their estimates $\hat{\sigma}^2$, $\hat{\nu}$, $\hat{\theta}$) in all our Examples 1–3.

6. 'Student' processes

The scaled t-distribution symmetric about 0 and with degrees of freedom $\nu > 0$ (not necessarily an integer) has PDF

$$f_T(x; \nu, \delta) = \frac{\Gamma((\nu+1)/2)}{\delta\sqrt{\pi}\,\Gamma(\nu/2)} \frac{1}{(1+(x/\delta)^2)^{(\nu+1)/2}}, \tag{36}$$

with $-\infty < x < \infty$, where $\delta > 0$ is a scaling parameter. The classical form ('Student's' t-distribution) has $\delta = \sqrt{n}$, $\nu = n$, with n a positive integer. A random variable X with PDF (36) has characteristic function:

$$\phi_T(u) = \mathrm{E}\,\mathrm{e}^{iuX} = \frac{(\delta|u|)^{\nu/2} K_{\nu/2}(\delta|u|)}{\Gamma(\nu/2)2^{\nu/2-1}}, \tag{37}$$

with $-\infty < u < \infty$. This form was given by Hurst *et al.* (1997, Equation (54)) and attributed to unpublished work of Hurst (1995). It was already available from Madan and Seneta (1990) by a simple duality argument. See also below for another justification. The point to be made for the present is that Praetz (1972) took the PDF (36) to describe the distribution of the symmetrically distributed mean-corrected increment (4), of i.i.d. increments. The distribution is thus sometimes called Praetz's t in financial contexts, and on account of its Pareto-type (power-law) tails has been the preferred distribution in modelling (see e.g. Hurst *et al.* (1997)). Heyde and Leonenko (2003) review it as the increment distribution in the setting of dependence structure of the log-price process.

The corresponding distribution of $\tau_t = T_t - T_{t-1}$ is given (in the setting of Section 1) by the reciprocal (inverse) gamma PDF

$$f_\tau(w) = \begin{cases} \dfrac{(\nu/2-1)^{\nu/2} w^{-\nu/2-1} \mathrm{e}^{-(\nu/2-1)/w}}{\Gamma(\nu/2)}, & w > 0, \\ 0 & \text{elsewhere,} \end{cases} \tag{38}$$

where $\mathrm{E}\,\tau_t = 1$ as required, providing $\nu > 2$. This gives the PDF of $\sigma(W(T_t) - W(T_{t-1}))$ as (36) with $\delta = \sigma(\nu-2)^{1/2}$. Heyde and Gay (2000) use the PDF (38) in conjunction with the model of (1) and (2), with $\theta = -\sigma^2/2$, to derive option-pricing formulae. Heyde and Leonenko (2003) show rigorously that a FATGBM model corresponding to this PDF exists providing $\nu > 4$ is an integer.

The PDFs (17) and (38) are extreme cases of the PDF form of a generalized inverse Gaussian distribution (Hurst *et al.* (1997, p. 107)). If V is described by a PDF of this form, then so is $1/V$. In general the distribution of τ_t can be taken as generalized inverse Gaussian. The relation

$$\mathrm{E}(\mathrm{e}^{\mathrm{i}u(W(T_t)-W(T_{t-1}))}) = \mathrm{E}(\mathrm{e}^{-u^2\tau_t/2}), \tag{39}$$

with u real for the characteristic function—since putting $s = u^2/2$ in (39) gives the Laplace transform of the distribution of τ_t on the right—shows that the distributions of $W(T_t) - W(T_{t-1})$ and τ_t determine each other uniquely, and provides a mechanism for producing the transform of one from the other. Nguyen *et al.* (2003) show positive skewness of generalized inverse Gaussian diatributions, so M_3 in (8) is positive in this case generally.

To show simply that (37) is the characteristic function corresponding to (36) with $\delta = (\nu-2)^{1/2}$ when $\nu > 2$, write an integral expression for the Laplace transform of (38), invoke the first expression for $K_\eta(\cdot)$ in Section 3, and then use (39).

The (symmetric) VG distribution was introduced as a direct competitor to Praetz's t—and even in the same journal—by Madan and Seneta (1990), by taking the mean-corrected log-price increment to have an $\mathcal{N}(0, \sigma^2 V)$ distribution, where V has a gamma distribution; whereas Praetz's t is $\mathcal{N}(0, \sigma^2 / V)$ distributed. The VG was seen, *inter alia,* to have a particularly simple characteristic function, (22) and (32), compared to that of the t-distribution, (37). However, the two distributions are dual to each other in several senses, and continuing study of the VG model at least in a manner complementary to Heyde and Kou (2002) and Heyde and Leonenko (2003), is envisaged to further elucidate its value as a model.

References

CARR, P. AND MADAN, D. B. (1999). Option valuation using the fast Fourier transform. *J. Comput. Finance* **1,** 61–73.

EPPS, T. W. (2000). *Pricing Derivative Securities.* World Scientific, Singapore.

ERDÉLYI, A., MAGNUS, W., OBERHETTINGER, F. AND TRICOMI, F. G. (1953). *Bateman Manuscript Project: Tables of Integral Transforms*, Vol. 2. McGraw-Hill, New York.

HEYDE, C. C. (1999). A risky asset model with strong dependence through fractal activity time. *J. Appl. Prob.* **36,** 1234–1239.

HEYDE, C. C. AND GAY, R. (2000). Fractals and contingent claims. Preprint.

HEYDE, C. C. AND KOU, S. G. (2002). On the controversy over tailweight of distributions. Preprint.

HEYDE, C. C. AND LEONENKO, N. N. (2003). Student and related processes. Preprint.

HEYDE, C. C. AND LIU, S. (2001). Empirical realities for a minimal description risky asset model. The need for fractal features. *J. Korean Math. Soc.* **38,** 1047–1059.

HURST, S. R. (1995). The characteristic function of the Student t distribution. Res. Rep. FMRR 006–95, Centre for Financial Mathematics, The Australian National University, Canberra.

HURST, S. R. AND PLATEN, E. (1997). The marginal distribution of returns and volatility. In *L_1-Statistical Procedures and Related Topics* (IMS Lecture Notes Monogr. Ser. **31**), ed. Y. Dodge, Institute for Mathematical Statistics, Hayward, CA, pp. 301–314.

HURST, S. R., PLATEN, E. AND RACHEV, S. R. (1997). Subordinated Markov models: a comparison. *Financial Eng. Japanese Markets* **4,** 97–124.

MADAN, D. B. AND MILNE, F. (1991). Option pricing with VG martingale components. *Math. Finance* **1,** 39–55.

MADAN, D. B. AND SENETA, E. (1987a). Characteristic function estimation using maximum likelihood on transformed variables. Econometrics Discussion Papers 87–08, Department of Econometrics, University of Sydney.

MADAN, D. B. AND SENETA, E. (1987b). Chebyshev polynomial approximations and characteristic function estimation. *J. R. Statist. Soc. B* **49,** 163–169.

MADAN, D. B. AND SENETA, E. (1987c). Simulation of estimates using the empirical characteristic function. *Internat. Statist. Rev.* **55,** 153–161.

MADAN, D. B. AND SENETA, E. (1989). Chebyshev polynomial approximation for characteristic function estimation. Some theoretical supplements. *J. R. Statist Soc. B* **51,** 281–285.

MADAN, D. B. AND SENETA, E. (1990). The variance-gamma (V. G.) model for share market returns. *J. Business* **63,** 511–524.

MADAN, D. B., CARR, P. P. AND CHANG, E. C. (1998). The variance gamma process and option pricing. *Europ. Finance Rev.* **2,** 79–105.

NGUYEN, T. T., CHEN, J. T., GUPTA, A. K. AND DINH, K. T. (2003). A proof of the conjecture on positive skewness of generalized inverse Gaussian distributions. *Biometrika* **90,** 245–250.

PRAETZ, P. D. (1972). The distribution of share price changes. *J. Business* **45,** 49–55

TJETJEP, A. A. (2002). The variance-gamma asset pricing models. Honours Project, School of Mathematics and Statistics, University of Sydney.

WHITTAKER, E. T. AND WATSON, G. N. (1915). *A Course of Modern Analysis.* Cambridge University Press.